

RECEIVED

APR 03 2002

TECH CENTER 1600/2900



1600

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/09/893,666A

TIME: 13:31:48

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\02282002\I893666A.raw

3 <110> APPLICANT: YAMASHITA, ICHIRO
 5 <120> TITLE OF INVENTION: High estrogen-sensitive medaka fish
 7 <130> FILE REFERENCE: 210217US-620-7249-0
 9 <140> CURRENT APPLICATION NUMBER: US 09/893,666A
 C--> 10 <141> CURRENT FILING DATE: 2002-02-12
 12 <150> PRIOR APPLICATION NUMBER: JP 2000-247729
 13 <151> PRIOR FILING DATE: 2000-08-17
 15 <160> NUMBER OF SEQ ID NOS: 7
 17 <170> SOFTWARE: PatentIn version 3.1
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 2764
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Oryzias latipes
 24 <220> FEATURE:
 25 <221> NAME/KEY: CDS
 26 <222> LOCATION: (211)..(1935)
 27 <223> OTHER INFORMATION:
 30 <400> SEQUENCE: 1
 31 gtctcgtctgc tagatgcctg tcaggcaggc agagaggaag cagcccgtgt tgcgcagcac 60
 33 atctgaggat gattcatgag taagagacag agctcgggtgc agatcaggca gctgttcgga 120
 35 ccagcactca gatccaggat cagcccagcc tcctcagagc tggagaccct cccccacct 180
 37 cgcctctcgc cccgtgaccc cctcgggtgac atg tac cct gaa gag agc cgg ggt 234
 38 Met Tyr Pro Glu Glu Ser Arg Gly
 39 1 5
 41 tct gga ggg gtg gct gct gtg gac ttt ttg gaa ggg acg tac gac tat 282
 42 Ser Gly Gly Val Ala Ala Val Asp Phe Leu Glu Gly Thr Tyr Asp Tyr
 43 10 15 20
 45 gcc gcc gcc aac cct gcc acg act ccc ctt tac agc cag tcc agc acc 330
 46 Ala Ala Pro Asn Pro Ala Thr Thr Pro Leu Tyr Ser Gln Ser Ser Thr
 47 25 30 35 40
 49 ggc tac tac tct gct ccc ctg gaa aca aac gga ccc ccc tca gaa ggc 378
 50 Gly Tyr Tyr Ser Ala Pro Leu Glu Thr Asn Gly Pro Pro Ser Glu Gly
 51 45 50 55
 53 agt ctg cag tcc ctg gcc agt ggg ccg acg agc cct ctg gtg ttt gtg 426
 54 Ser Leu Gln Ser Leu Gly Ser Gly Pro Thr Ser Pro Leu Val Phe Val
 55 60 65 70
 57 ccc tcc agc ccc aga ctc agt ccc ttt atg cat cca ccc agc cac cac 474
 58 Pro Ser Ser Pro Arg Leu Ser Pro Phe Met His Pro Pro Ser His His
 59 75 80 85
 61 tat ctg gaa acc act tcc acg ccc gtt tac aga tcc agc cac cag gga 522
 62 Tyr Leu Glu Thr Thr Ser Thr Pro Val Tyr Arg Ser Ser His Gln Gly
 63 90 95 100
 65 gcc tcc agg gag gac cag tgc gcc tcc cgg gag gac acg tgc agc ctg 570

RECEIVED
 MAR 1 2 2002
 TECH CENTER 1600/2900

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/09/893,666A

TIME: 13:31:48

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\02282002\I893666A.raw

```

66 Ala Ser Arg Glu Asp Gln Cys Gly Ser Arg Glu Asp Thr Cys Ser Leu
67 105 110 115 120
69 ggg gag tta ggc gcc gga gcc ggg gct ggg ggg ttt gag atg gcc aaa 618
70 Gly Glu Leu Gly Ala Gly Ala Gly Ala Gly Gly Phe Glu Met Ala Lys
71 125 130 135
73 gac acg cgt ttc tgc gcc gtg tgc agc gac tac gcc tct ggg tac cac 666
74 Asp Thr Arg Phe Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His
75 140 145 150
77 tat ggg gtg tgg tct tgt gag ggc tgc aag gcc ttc ttc aag agg agc 714
78 Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser
79 155 160 165
81 atc cag ggt cac aat gac tat atg tgc cca gcg acc aat cag tgc act 762
82 Ile Gln Gly His Asn Asp Tyr Met Cys Pro Ala Thr Asn Gln Cys Thr
83 170 175 180
85 att gac aga aat cgg agg aag agc tgc cag gct tgt cgt ctt agg aag 810
86 Ile Asp Arg Asn Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys
87 185 190 195 200
89 tgt tac gaa gtg gga atg atg aaa ggc ggt gtg cgc aag gac cgc att 858
90 Cys Tyr Glu Val Gly Met Met Lys Gly Gly Val Arg Lys Asp Arg Ile
91 205 210 215
93 cgc att tta cgg cgt gac aaa cgg cgg aca ggc gtt ggt gat gga gac 906
94 Arg Ile Leu Arg Arg Asp Lys Arg Arg Thr Gly Val Gly Asp Gly Asp
95 220 225 230
97 aag gtt gta aag ggt cag gag cat aaa acg gtg cat tat gat gga agg 954
98 Lys Val Val Lys Gly Gln Glu His Lys Thr Val His Tyr Asp Gly Arg
99 235 240 245
101 aaa cgc agc agc aca gga gga gga gga gga gga gga gga gga aga ctg 1002
102 Lys Arg Ser Ser Thr Gly Gly Gly Gly Gly Gly Gly Gly Gly Arg Leu
103 250 255 260
105 tct gtg acc agc ata cct cct gag cag gtg ctg ctc ctc ctt cag ggc 1050
106 Ser Val Thr Ser Ile Pro Pro Glu Gln Val Leu Leu Leu Leu Gln Gly
107 265 270 275 280
109 gcc gag ccc ccg ata ctc tgc tcg cgt cag aag ttg agc cga ccg tac 1098
110 Ala Glu Pro Pro Ile Leu Cys Ser Arg Gln Lys Leu Ser Arg Pro Tyr
111 285 290 295
113 acc gag gtc acc atg atg acc ctg ctc acc agc atg gca gac aag gag 1146
114 Thr Glu Val Thr Met Met Thr Leu Leu Thr Ser Met Ala Asp Lys Glu
115 300 305 310
117 ctg gtc cac atg atc gcc tgg gcc aag aag ctc cca ggt ttt ctg cag 1194
118 Leu Val His Met Ile Ala Trp Ala Lys Lys Leu Pro Gly Phe Leu Gln
119 315 320 325
121 ctg tcc ctg cac gat cag gtg ctg ctg ctg gag agc tcg tgg ctg gag 1242
122 Leu Ser Leu His Asp Gln Val Leu Leu Leu Glu Ser Ser Trp Leu Glu
123 330 335 340
125 gtg ctc atg atc ggc ctc att tgg agg tcc atc cac tgt ccc ggg aag 1290
126 Val Leu Met Ile Gly Leu Ile Trp Arg Ser Ile His Cys Pro Gly Lys
127 345 350 355 360
129 ctc atc ttt gca caa gac ctc atc ctg gac agg aat gag gga gac tgc 1338
130 Leu Ile Phe Ala Gln Asp Leu Ile Leu Asp Arg Asn Glu Gly Asp Cys

```

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/09/893,666A

TIME: 13:31:48

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\02282002\I893666A.raw

131		365		370		375	
133	gtg gaa ggc atg acg gag atc ttc gac atg ctg ctg gcc act gct tcc						1386
134	Val Glu Gly Met Thr Glu Ile Phe Asp Met Leu Leu Ala Thr Ala Ser						
135		380		385		390	
137	cgc ttc cgt gtg ctc aaa ctc aaa cct gag gaa ttc gtc tgc ctc aaa						1434
138	Arg Phe Arg Val Leu Lys Leu Lys Pro Glu Glu Phe Val Cys Leu Lys						
139		395		400		405	
141	gct att att tta ctc aac tcc ggt gct ttt tct ttc tgc acc ggc acc						1482
142	Ala Ile Ile Leu Leu Asn Ser Gly Ala Phe Ser Phe Cys Thr Gly Thr						
143		410		415		420	
145	atg gag cca ctt cac aac agc gcg gcg gtt cag agc atg ctg gac acc						1530
146	Met Glu Pro Leu His Asn Ser Ala Ala Val Gln Ser Met Leu Asp Thr						
147	425		430		435		440
149	atc aca gac gca ctc att cat tac atc agt cag tcg ggt tac ttg gcc						1578
150	Ile Thr Asp Ala Leu Ile His Tyr Ile Ser Gln Ser Gly Tyr Leu Ala						
151		445		450		455	
153	cag gag cag gcg aga cgg cag gcc cag ctg ctc ctg ctg ctc tcc cac						1626
154	Gln Glu Gln Ala Arg Arg Gln Ala Gln Leu Leu Leu Leu Ser His						
155		460		465		470	
157	atc agg cac atg agc aac aaa ggc atg gag cac ctc tac agc atg aag						1674
158	Ile Arg His Met Ser Asn Lys Gly Met Glu His Leu Tyr Ser Met Lys						
159		475		480		485	
161	tgc aag aac aaa gtc cct ctt tat gac ctc cta ctg gag atg ctc gat						1722
162	Cys Lys Asn Lys Val Pro Leu Tyr Asp Leu Leu Leu Glu Met Leu Asp						
163		490		495		500	
165	gcc cac cgc ctg cac cac ccc gtc aga gca ccc cag tcc ttg tcc caa						1770
166	Ala His Arg Leu His His Pro Val Arg Ala Pro Gln Ser Leu Ser Gln						
167	505		510		515		520
169	gtc gac aga gac cct ccc tcc acc agc agc ggc ggg ggt gga atc gct						1818
170	Val Asp Arg Asp Pro Pro Ser Thr Ser Ser Gly Gly Gly Gly Ile Ala						
171		525		530		535	
173	ccc ggt tct ata tca gca tct cga ggc aga atc gag agt ccg agc aga						1866
174	Pro Gly Ser Ile Ser Ala Ser Arg Gly Arg Ile Glu Ser Pro Ser Arg						
175		540		545		550	
177	ggc ccc ttt gct ccc agt gtc ctt cag tat gga ggg tcg cgt cct gac						1914
178	Gly Pro Phe Ala Pro Ser Val Leu Gln Tyr Gly Gly Ser Arg Pro Asp						
179		555		560		565	
181	tgc acc ccg gcc ctt caa gac tgagcacaca gtccaaggcc ctttttttgt						1965
182	Cys Thr Pro Ala Leu Gln Asp						
183		570		575			
185	ggctcaaggg ttcaggttg gacaagggtga tgcttgattt aattttaaga attatttata						2025
187	aataagagtg gcgctgagag gagaagctcc cacaatgaac tgctctgct tggctcagct						2085
189	tttgtgcagt cactttaatc tgcttatatt catctccttt gtaaacctga gcgtctcttt						2145
191	agcagctttt ttttgctctc caaacagcat gtggtagatt gtaagggttc gtcccatgag						2205
193	ttctggtgat ttcaagaaaa tgagcagcta atgttttctg taaccgtctt gacccaagtg						2265
195	cacttctctc tggattaaag gggctaattg gcattatttt gtctcttgta catatgggat						2325
197	ggctaagaat aatgagagta attgtcagat ttgtgtaga acttaccac aaatgaatt						2385
199	ttaaaataag atttaaaaac aaaagaggca agatcaaacc tgagagcaact gaagacacgc						2445
201	tgtagaaagc tgggtaaatt tgttatccac gtctatctct ggaaaggact ttgttctctg						2505

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/09/893,666A

TIME: 13:31:48

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\02282002\I893666A.raw

```

203 tgcctgcagc tcatttactc tgaacttgct acttggtgaa catttggtgca cttgtccgtg 2565
205 tttttctagc actgtagctt atgaacgctg agaaagaatc taatgctttg atgcacagat 2625
207 ttgccttgta ttgtacatct cagccacaaa cgtacttttc gtccacaagt tgactgactg 2685
209 caccttgatt aaattgtcta aaagtgcatt taaatgttga attctgtgaa aattaaaaag 2745
211 gcaattcctg tttctatatt 2764
214 <210> SEQ ID NO: 2
215 <211> LENGTH: 575
216 <212> TYPE: PRT
217 <213> ORGANISM: Oryzias latipes
219 <400> SEQUENCE: 2
221 Met Tyr Pro Glu Ser Arg Gly Ser Gly Gly Val Ala Ala Val Asp
222 1 5 10 15
225 Phe Leu Glu Gly Thr Tyr Asp Tyr Ala Ala Pro Asn Pro Ala Thr Thr
226 20 25 30
229 Pro Leu Tyr Ser Gln Ser Ser Thr Gly Tyr Tyr Ser Ala Pro Leu Glu
230 35 40 45
233 Thr Asn Gly Pro Pro Ser Glu Gly Ser Leu Gln Ser Leu Gly Ser Gly
234 50 55 60
237 Pro Thr Ser Pro Leu Val Phe Val Pro Ser Ser Pro Arg Leu Ser Pro
238 65 70 75 80
241 Phe Met His Pro Pro Ser His His Tyr Leu Glu Thr Thr Ser Thr Pro
242 85 90 95
245 Val Tyr Arg Ser Ser His Gln Gly Ala Ser Arg Glu Asp Gln Cys Gly
246 100 105 110
249 Ser Arg Glu Asp Thr Cys Ser Leu Gly Glu Leu Gly Ala Gly Ala Gly
250 115 120 125
253 Ala Gly Gly Phe Glu Met Ala Lys Asp Thr Arg Phe Cys Ala Val Cys
254 130 135 140
257 Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly
258 145 150 155 160
261 Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Met
262 165 170 175
265 Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Arg Asn Arg Arg Lys Ser
266 180 185 190
269 Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Met Lys
270 195 200 205
273 Gly Gly Val Arg Lys Asp Arg Ile Arg Ile Leu Arg Arg Asp Lys Arg
274 210 215 220
277 Arg Thr Gly Val Gly Asp Gly Asp Lys Val Val Lys Gly Gln Glu His
278 225 230 235 240
281 Lys Thr Val His Tyr Asp Gly Arg Lys Arg Ser Ser Thr Gly Gly Gly
282 245 250 255
285 Gly Gly Gly Gly Gly Gly Arg Leu Ser Val Thr Ser Ile Pro Pro Glu
286 260 265 270
289 Gln Val Leu Leu Leu Leu Gln Gly Ala Glu Pro Pro Ile Leu Cys Ser
290 275 280 285
293 Arg Gln Lys Leu Ser Arg Pro Tyr Thr Glu Val Thr Met Met Thr Leu
294 290 295 300
297 Leu Thr Ser Met Ala Asp Lys Glu Leu Val His Met Ile Ala Trp Ala

```

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/09/893,666A

TIME: 13:31:48

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\02282002\I893666A.raw

```

298 305          310          315          320
301 Lys Lys Leu Pro Gly Phe Leu Gln Leu Ser Leu His Asp Gln Val Leu
302          325          330          335
305 Leu Leu Glu Ser Ser Trp Leu Glu Val Leu Met Ile Gly Leu Ile Trp
306          340          345          350
309 Arg Ser Ile His Cys Pro Gly Lys Leu Ile Phe Ala Gln Asp Leu Ile
310          355          360          365
313 Leu Asp Arg Asn Glu Gly Asp Cys Val Glu Gly Met Thr Glu Ile Phe
314          370          375          380
317 Asp Met Leu Leu Ala Thr Ala Ser Arg Phe Arg Val Leu Lys Leu Lys
318 385          390          395          400
321 Pro Glu Glu Phe Val Cys Leu Lys Ala Ile Ile Leu Leu Asn Ser Gly
322          405          410          415
325 Ala Phe Ser Phe Cys Thr Gly Thr Met Glu Pro Leu His Asn Ser Ala
326          420          425          430
329 Ala Val Gln Ser Met Leu Asp Thr Ile Thr Asp Ala Leu Ile His Tyr
330          435          440          445
333 Ile Ser Gln Ser Gly Tyr Leu Ala Gln Glu Gln Ala Arg Arg Gln Ala
334          450          455          460
337 Gln Leu Leu Leu Leu Leu Ser His Ile Arg His Met Ser Asn Lys Gly
338 465          470          475          480
341 Met Glu His Leu Tyr Ser Met Lys Cys Lys Asn Lys Val Pro Leu Tyr
342          485          490          495
345 Asp Leu Leu Leu Glu Met Leu Asp Ala His Arg Leu His His Pro Val
346          500          505          510
349 Arg Ala Pro Gln Ser Leu Ser Gln Val Asp Arg Asp Pro Pro Ser Thr
350          515          520          525
353 Ser Ser Gly Gly Gly Gly Ile Ala Pro Gly Ser Ile Ser Ala Ser Arg
354          530          535          540
357 Gly Arg Ile Glu Ser Pro Ser Arg Gly Pro Phe Ala Pro Ser Val Leu
358 545          550          555          560
361 Gln Tyr Gly Gly Ser Arg Pro Asp Cys Thr Pro Ala Leu Gln Asp
362          565          570          575

```

365 <210> SEQ ID NO: 3

366 <211> LENGTH: 20

367 <212> TYPE: DNA

368 <213> ORGANISM: ARTIFICIAL SEQUENCE

370 <220> FEATURE:

371 <223> OTHER INFORMATION: SYNTHETIC DNA

373 <400> SEQUENCE: 3

374 tcggtgacat gtaccctgaa

20

377 <210> SEQ ID NO: 4

378 <211> LENGTH: 20

379 <212> TYPE: DNA

380 <213> ORGANISM: ARTIFICIAL SEQUENCE

382 <220> FEATURE:

383 <223> OTHER INFORMATION: SYNTHETIC DNA

385 <400> SEQUENCE: 4

386 ctgtgtgctc agtcttgaag

20

VERIFICATION SUMMARY

DATE: 02/28/2002

PATENT APPLICATION: US/09/893,666A

TIME: 13:31:49

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\02282002\I893666A.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date